fullness at t_i^e (b_i^e), plus the number of bits entering the decoder VBV buffer between t_i^e and

$$t_i\left(\int_{t_i^e}^{t_i^e+L} R_{vol,decoder}(t)\,dt\right),$$

minus the number of bits removed form the decoder buffer between t_i^e and t_i. The number of bits removed is the sum of encoder buffer occupancy at t_i^e immediately before adding VOP i (eb_i^e) and decoder buffer occupancy at t_i^e (b_i^e) because all bitstream data prior to VOP i must be consumed before VOP i can be decoded. The later two quantities represent bitstream data prior to VOP i since VOP i has not been added to the encoder buffer. Therefore, total bits in the decoder buffer are bounded by

$$b_i^e + \int_{t_i^e}^{t_i^e+L} R_{vol,decoder}(t)\,dt - (eb_i^e + b_i^e) < B$$

Which yields

$$\int_{t_i^e}^{t_i^e+L} R_{vol,decoder}(t)\,dt - eb_i^e + b_i^e < B.$$

Therefore, encoder_buffer_fullness at t_i^e is lower bounded by

$$eb_i^e > \int_{t_i^e}^{t_i^e+L} R_{vol,decoder}(t)\,dt - B = T_1.$$

The same arguments given above about the constant delay channel can be applied here. Also, for the type of channels which have known maximum transmission rate $R_{vol,max}$, T_1 can be set to be a upper bound of

$$\int_{t_i^e}^{t_i^e+L} R_{vol,decoder}(t)\,dt - B \text{ as } T_1 = L \cdot R_{vol,max} - B.$$

The bounds T_1 and T_2 are checked in the rate-control algorithm and the corrective action are to perform bit-allocation of the VOPs and adjust quantization levels of the coding units (e.g. VOP, macroblocks).

The encoder must take the following corrective action if the simulated decoder VBV buffer gets too full or too empty:

1. If the simulated decoder VBV buffer becomes too full (i.e. the encoder VBV buffer is too empty), the encoder can correct the problem by:
   (a) reducing the quantization level to generate large VOPs, or
   (b) outputting stuffing bits at the end of the VOP. Note that generating larger VOPs reduces the decoder VBV occupancy.
2. If the simulated decoder VBV becomes too empty (i.e. the encoder VBV buffer is too full), the encoder can correct the problem by:
   (a) increasing quantization levels to generate fewer bits, or
   (b) delaying the generation of the next VOP (often called skipping VOP), or
   (c) zeroing the high frequency coefficients to reduce the number of bits/VOP generated.

It should now be appreciated that the present invention provides a video rate buffer model for bounding the memory requirements of a video decoder in a push dataflow scenario. The rate buffer model of the present invention constrains the video encoder to producing bitstreams that are decodable with a predetermined buffer memory size. Thus, push dataflow applications are efficiently accommodated.

What is claimed is:

1. An encoder apparatus for enabling a push dataflow bitstream without causing a modeled data buffer of predetermined memory size for the pushed data at a decoder to overflow or underflow, comprising:

a processor adapted to encode data to provide the bitstream for communication to a decoder; wherein:

the encoded data comprises at least one video or visual object (VO) with at least one video or visual object layer (VOL), including an associated header, followed by at least one video or visual object plane (VOP);

a field in the VOL header designates an occupancy of the buffer just before removal of the first VOP following the VOL header from the buffer; and

said processor uses a simulation to simulate the decoder buffer and controls the bitstream in response to the simulation to preclude overflow or underflow of the decoder buffer.

2. The apparatus of claim 1, wherein:

when the buffer is initially empty, the occupancy field is examined to determine an initial occupancy of the buffer before decoding the initial VOP.

3. The apparatus of claim 1, wherein:

the processor provides a flag to control the inclusion of at least one field in the VOL header when equivalent information is not present in an encapsulating system multiplex.

4. The apparatus of claim 3, wherein:

the flag allows a visual elementary stream of the bitstream as a standalone entity to specify a buffer model.

5. The apparatus of claim 3, wherein:

the at least one field whose inclusion in the VOL header is controlled by the flag designates a VOP rate of the bitstream.

6. The apparatus of claim 3, wherein:

the at least one field whose inclusion in the VOL header is controlled by the flag designates a peak bit rate of the bitstream.

7. The apparatus of claim 3, wherein:

the at least one field whose inclusion in the VOL header is controlled by the flag designates whether the VOL contains at least one B-VOP.

8. The apparatus of claim 3, wherein:

the at least one field whose inclusion in the VOL header is controlled by the flag designates the size of the modeled buffer.

9. The apparatus of claim 3, wherein:

the at least one field whose inclusion in the VOL header is controlled by the flag designates said field that designates the occupancy of the buffer.

10. The apparatus of claim 1, wherein:

the video or visual object comprises a video object.

11. The apparatus of claim 1, wherein:

the video or visual object comprises a still texture object.

12. The apparatus of claim 1, wherein:

the video or visual object comprises a mesh object.

13. The apparatus of claim 1, wherein:

the video or visual object comprises a face object.

14. The apparatus of claim 1, wherein:

the encoded data comprises a plurality of VOLs, and a decoder buffer model is applied independently to each VOL using buffer size and rate functions particular to each VOL.

15. The apparatus of claim 1, wherein:

the bitstream is compatible with an MPEG-4 video coding standard.

16. The apparatus of claim 1, wherein:

additional fields are provided in respective subsequent VOL headers to designate respective subsequent buffer occupancy levels.

17. The apparatus of claim 16, wherein:

the processor maintains a difference between the additional fields in the subsequent VOL headers and a running cumulative buffer occupancy just before removal of a VOP from the buffer within a tolerance.

18. The apparatus of claim 1, wherein:

the processor determines a size (d_i) of a current VOP as a number of bits extending to the last bit of the current VOP and starting from either the last bit of the previous VO or the first bit of a start code for the first VOP of the encoded data.

19. The apparatus of claim 1, wherein:

the processor determines a decoding time t_i of an ith VOP, wherein:

(a) $t_i=\tau_i$ if the VOL contains no B-VOPs, where τ_i is a composition time of the ith VOP, and

(b) $t_i=\tau_i-m_i$, when the ith VOP is an anchor VOP, and m_i accounts for a delay of at least one immediately subsequent B-VOP to be composited or presented.

20. The apparatus of claim 1, wherein:

the processor determines a decoding time t_i of an ith VOP, wherein:

(a) $t_i=\tau_i$ if the VOL contains no B-VOPs, where τ_i is a presentation time of the ith VOP when the decoder is a no-compositor decoder, and

(b) $t_i=\tau_i-m_i$, when the ith VOP is an anchor VOP, and m_i accounts for a delay of at least one immediately subsequent B-VOP to be composited or presented.

21. The apparatus of claim 1, wherein:

said pushed data flow comprises video data that includes intraframe (I), prediction (P) and bi-directional (B) video object planes (VOPs).

22. The apparatus of claim 21, wherein said processor controls said bitstream by at least one of:

allocating bits between different VOPs, and

adjusting quantization levels of coding units forming said VOPs.

23. The apparatus of claim 22, wherein:

the modeled data buffer comprises a visual or video buffering verifier (VBV) buffer.

24. The apparatus of claim 23, wherein:

said processor monitors the VBV buffer and, when the simulation indicates that the VBV buffer has or will become too full, a quantization level for said coding units is reduced.

25. The apparatus of claim 23, wherein:

said processor monitors the VBV buffer and, when the simulation indicates that the VBV buffer has or will become too empty, said quantization level for said coding units is increased.

26. The apparatus of claim 23, wherein:

said processor monitors the VBV buffer and, when the simulation indicates that the VBV buffer has or will become too empty, the generation of the next VOP is delayed.

27. The apparatus of claim 23, wherein:

said processor monitors the VBV buffer and, when the simulation indicates that the VBV buffer has or will become too empty, high frequency coefficients of said coding units are zeroed to reduce the number of bits generated per VOP.

28. The apparatus of claim 23, wherein:

said processor monitors the VBV buffer and, when it is determined that the VBV buffer has or will become too full, stuffing bits are added to the end of at least one VOP.

29. The apparatus of claim 1, wherein:

the processor includes an encoder buffer for receiving the encoded data prior to providing the bitstream therefrom; and

the processor controls the rate of the bitstream such that the fullness of the encoder buffer eb_i^e after encoding an ith VOP is upper bounded by

$$\int_{t_i^e}^{t_i^e+L} R_{vol,decoder}(t)\,dt - d_i,$$

where t_i^e is the time of starting to encode the ith VOP, L is the time difference between the encoding time t_i^e and the decoding time of the ith VOP, d_i is the amount of coded data for the ith VOP, and $R_{vol,decoder}(t)$ is the instantaneous channel bit rate seen by the decoder.

30. The apparatus of claim 1, wherein:

the processor includes an encoder buffer for receiving the encoded data prior to providing the bitstream therefrom; and

the processor controls the rate of the bitstream such that the fullness of the encoder buffer eb_i^e after encoding an ith VOP is upper bounded by

$$\int_{t_i^e-L}^{t_i^e} R_{vol,encoder}(t)\,dt - d_i,$$

where t_i^e is the time of starting to encode the ith VOP, L is the time difference between the encoding time t_i^e and the decoding time of the ith VOP, d_i is the amount of coded data for the ith VOP, and $R_{vol,encoder}(t)$ is the instantaneous channel bit rate seen by the encoder.

31. The apparatus of claim 1, wherein:

the processor includes an encoder buffer for receiving the encoded data prior to providing the bitstream therefrom; and

the processor controls the rate of the bitstream such that the fullness of the encoder buffer eb_i^e after encoding an ith VOP is lower bounded by

$$\int_{t_i^e}^{t_i^e+L} R_{vol,decoder}(t)\,dt - B,$$

where t_i^e is the time of starting to encode the ith VOP, L is the time difference between the encoding time t_i^e and the decoding time of the ith VOP, B is the size of the decoder buffer, and $R_{vol,decoder}(t)$ is the instantaneous channel bit rate seen by the decoder.

32. The apparatus of claim 1, wherein:

the processor includes an encoder buffer for receiving the encoded data prior to providing the bitstream there-

from: and the processor controls the rate of the bitstream such that the fullness of the encoder buffer eb_i^e after encoding an ith VOP is lower bounded by

$$\int_{t_i^e - L}^{t_i^e} R_{vol,encoder}(t)\,dt - B,$$

where t_i^e is the time of starting to encode the ith VOP, L is the time difference between the encoding time t_i^e and the decoding time of the ith VOP, B is the size of the decoder buffer, and $R_{vol,encoder}(t)$ is the instantaneous channel bit rate seen by the encoder.

33. An encoding method for enabling a push dataflow bitstream without causing a modeled data buffer of predetermined memory size for the pushed data at a decoder to overflow or underflow, comprising the steps of:

encoding data to provide the bitstream for communication to a decoder; wherein:

the encoded data comprises at least one video or visual object (VO) with at least one video or visual object layer (VOL), including an associated header, followed by at least one video or visual object plane (VOP); and

a field in the VOL header designates an occupancy of the buffer just before removal of the first VOP following the VOL header from the buffer; and

using a simulation to simulate the decoder buffer and control the bitstream in response to the simulation to preclude overflow or underflow of the decoder buffer.

34. The method of claim 33, wherein:

when the buffer is initially empty, the occupancy field is examined to determine an initial occupancy of the buffer before decoding the initial VOP.

35. The method of claim 33, comprising the further step of:

providing a flag to control the inclusion of at least one field in the VOL header when equivalent information is not present in an encapsulating system multiplex.

36. The method of claim 35, wherein:

the flag allows a visual elementary stream of the bitstream as a standalone entity to specify a buffer model.

37. The method of claim 35, wherein:

the at least one field whose inclusion in the VOL header is controlled by the flag designates a VOP rate of the bitstream.

38. The method of claim 35, wherein:

the at least one field whose inclusion in the VOL header is controlled by the flag designates a peak bit rate of the bitstream.

39. The method of claim 35, wherein:

the at least one field whose inclusion in the VOL header is controlled by the flag designates whether the VOL contains at least one B-VOP.

40. The method of claim 35, wherein:

the at least one field whose inclusion in the VOL header is controlled by the flag designates the size of the modeled buffer.

41. The method of claim 35, wherein:

the at least one field whose inclusion in the VOL header is controlled by the flag designates said field that designates the occupancy of the buffer.

42. The method of claim 33, wherein:

the video or visual object comprises a video object.

43. The method of claim 33, wherein:

the video or visual object comprises a still texture object.

44. The method of claim 33, wherein:

the video or visual object comprises a mesh object.

45. The method of claim 33, wherein:

the video or visual object comprises a face object.

46. The method of claim 33, wherein:

the encoded data comprises a plurality of VOLs, and a decoder buffer model is applied independently to each VOL using buffer size and rate functions particular to each VOL.

47. The method of claim 33, wherein:

the bitstream is compatible with an MPEG-4 video coding standard.

48. The method of claim 33, wherein:

additional fields are provided in respective subsequent VOL headers to designate respective subsequent buffer occupancy levels.

49. The method of claim 48, comprising the further step of:

maintaining a difference between the additional fields in the subsequent VOL headers and a running cumulative buffer occupancy just before removal of a VOP from the buffer within a tolerance.

50. The method of claim 33, comprising the further step of:

determining a size (d_i) of a current VOP as a number of bits extending to the last bit of the current VOP and starting from either the last bit of the previous VO or the first bit of a start code for the first VOP of the encoded data.

51. The method of claim 33, comprising the further step of:

determining a decoding time t_i of an ith VOP, wherein:

(a) $t_i=\tau_i$ if the VOL contains no B-VOPs, where τ_i is a composition time of the ith VOP, and

(b) $t_i=\tau_i-m_i$, when the ith VOP is an anchor VOP, and m_i accounts for a delay of at least one immediately subsequent B-VOP to be composited or presented.

52. The method of claim 33, comprising the further step of:

determining a decoding time t_i of an ith VOP, wherein:

(a) $t_i=\tau_i$ if the VOL contains no B-VOPs, where τ_i is a presentation time of the ith VOP when the decoder is a no-compositor decoder, and

(b) $t_i=\tau_i-m_i$, when the ith VOP is an anchor VOP, and m_i accounts for a delay of at least one immediately subsequent B-VOP to be composited or presented.

53. The method of claim 33, wherein:

said pushed data flow comprises video data that includes intraframe (I), prediction (P) and bi-directional (B) video object planes (VOPs).

54. The method of claim 53, comprising the further step of controlling said bitstream by at least one of:

allocating bits between different VOPs, and

adjusting quantization levels of coding units forming said VOPs.

55. The method of claim 54, wherein:

the modeled data buffer comprises a visual or video buffering verifier (VBV) buffer.

56. The method of claim 55, comprising the further step of:

monitoring the VBV buffer and, when the simulation indicates that the VBV buffer has or will become too full, reducing a quantization level for said coding units.

57. The method of claim 55, comprising the further step of:

monitoring the VBV buffer and, when the simulation indicates that the VBV buffer has or will become too empty, increasing a quantization level for said coding units.

58. The method of claim 55, comprising the further step of:

monitoring the VBV buffer and, when the simulation indicates that the VBV buffer has or will become too empty, delaying the generation of the next VOP.

59. The method of claim 55, comprising the further step of:

monitoring the VBV buffer and, when the simulation indicates that the VBV buffer has or will become too empty, zeroing high frequency coefficients of said coding units to reduce the number of bits generated per VOP.

60. The method of claim 55, comprising the further step of:

monitoring the VBV buffer and, when it is determined that the VBV buffer has or will become too full, adding stuffing bits to the end of at least one VOP.

61. The method of claim 33, wherein an encoder buffer receives the encoded data prior to providing the bitstream therefrom, comprising the further step of:

controlling the rate of the bitstream such that the fullness of the encoder buffer eb_i^e after encoding an ith VOP is upper bounded by

$$\int_{t_i^e}^{t_i^e+L} R_{vol,decoder}(t)\,dt - d_i,$$

where t_i^e is the time of starting to encode the ith VOP, L is the time difference between the encoding time t_i^e and the decoding time of the ith VOP, d_i is the amount of coded data for the ith VOP, and $R_{vol,decoder}(t)$ is the instantaneous channel bit rate seen by the decoder.

62. The method of claim 33, wherein an encoder buffer receives the encoded data prior to providing the bitstream therefrom, comprising the further step of:

controlling the rate of the bitstream such that the fullness of the encoder buffer eb_i^e after encoding an ith VOP is upper bounded by

$$\int_{t_i^e-L}^{t_i^e} R_{vol,encoder}(t)\,dt - d_i,$$

where t_i^e is the time of starting to encode the ith VOP, L is the time difference between the encoding time t_i^e and the decoding time of the ith VOP, d_i is the amount of coded data for the ith VOP, and $R_{vol,encoder}(t)$ is the instantaneous channel bit rate seen by the encoder.

63. The method of claim 33, wherein an encoder buffer receives the encoded data prior to providing the bitstream therefrom, comprising the further step of:

controlling the rate of the bitstream such that the fullness of the encoder buffer eb_i^e after encoding an ith VOP is lower bounded by

$$\int_{t_i^e}^{t_i^e+L} R_{vol,decoder}(t)\,dt - B,$$

where t_i^e is the time of starting to encode the ith VOP, L is the time difference between the encoding time t_i^e and the decoding time of the ith VOP, B is the size of the decoder buffer, and $R_{vol,decoder}(t)$ is the instantaneous channel bit rate seen by the decoder.

64. The method of claim 33, wherein an encoder buffer receives the encoded data prior to providing the bitstream therefrom, comprising the further step of:

controlling the rate of the bitstream such that the fullness of the encoder buffer eb_i^e after encoding an ith VOP is lower bounded by

$$\int_{t_i^e-L}^{t_i^e} R_{vol,encoder}(t)\,dt - B,$$

where t_i^e is the time of starting to encode the ith VOP, L is the time difference between the encoding time t_i^e and the decoding time of the ith VOP, B is the size of the decoder buffer, and $R_{vol,encoder}(t)$ is the instantaneous channel bit rate seen by the encoder.

65. A decoder apparatus, comprising:

a data buffer of predetermined memory size; and

means for receiving a push dataflow bitstream that is obtained by encoding data in accordance with a model of the buffer so that the modeled buffer does not overflow or underflow; wherein:

the encoded data comprises at least one video or visual object (VO) with at least one video or visual object layer (VOL), including an associated header, followed by at least one video or visual object plane (VOP); and

a field in the VOL header designates an occupancy of the modeled buffer just before removal of the first VOP following the VOL header from the modeled buffer.

66. The apparatus of claim 65, wherein:

the encoded data comprises a flag that controls the inclusion of at least one field in the VOL header when equivalent information is not present in an encapsulating system multiplex.

67. The apparatus of claim 66, wherein:

the flag allows a visual elementary stream of the bitstream as a standalone entity to specify a buffer model.

68. The apparatus of claim 66, wherein:

the at least one field whose inclusion in the VOL header is controlled by the flag designates a VOP rate of the bitstream.

69. The apparatus of claim 66, wherein:

the at least one field whose inclusion in the VOL header is controlled by the flag designates a peak bit rate of the bitstream.

70. The apparatus of claim 66, wherein:

the at least one field whose inclusion in the VOL header is controlled by the flag designates whether the VOL contains at least one B-VOP.

71. The apparatus of claim 66, wherein:

the at least one field whose inclusion in the VOL header is controlled by the flag designates the size of the modeled buffer.

72. The apparatus of claim 66, wherein:

the at least one field whose inclusion in the VOL header is controlled by the flag designates said field that designates the occupancy of the buffer.

73. The apparatus of claim 65, wherein:

the video or visual object comprises a video object.

74. The apparatus of claim 65, wherein:

the video or visual object comprises a still texture object.

75. The apparatus of claim 65, wherein:

the video or visual object comprises a mesh object.

76. The apparatus of claim 65, wherein:

the video or visual object comprises a face object.

77. The apparatus of claim 65, wherein:

the encoded data comprises a plurality of VOLs, and a decoder buffer model is applied independently to each VOL using buffer size and rate functions particular to each VOL.

78. The apparatus of claim 65, wherein:

the bitstream is compatible with an MPEG-4 video coding standard.

79. The apparatus of claim 65, wherein:

additional fields are provided in respective subsequent VOL headers to designate respective subsequent buffer occupancy levels.

80. The apparatus of claim 65, wherein:

said pushed data flow comprises video data that includes intraframe (I), prediction (P) and bi-directional (B) video object planes (VOPs).

81. A method for providing a push dataflow bitstream at a decoder, comprising the steps of:

providing a data buffer of predetermined memory size at the decoder; and

receiving the push dataflow bitstream at the decoder, wherein said bitstream is obtained by encoding data in accordance with a model of the buffer so that the modeled buffer does not overflow or underflow; wherein:

the encoded data comprises at least one video or visual object (VO) with at least one video or visual object layer (VOL), including an associated header, followed by at least one video or visual object plane (VOP); and

a field in the VOL header designates an occupancy of the modeled buffer just before removal of the first VOP following the VOL header from the modeled buffer.

* * * * *

82. A bitstream for push dataflow communication to a decoder having a data buffer of predetermined memory size, said bitstream being obtained by encoding data in accordance with a model of the decoder buffer so that the modeled buffer does not overflow or underflow; wherein:

the encoded data of the bitstream comprises at least one video or visual object (VO) with at least one video or visual object layer (VOL), including an associated header, followed by at least one video or visual object plane (VOP); and

a field in the VOL header designates an occupancy of the modeled buffer just before removal of the first VOP following the VOL header from the modeled buffer.

83. The bitstream of claim 82, wherein:

the encoded data further comprises a flag that controls the inclusion of at least one field in the VOL header when equivalent information is not present in an encapsulating system multiplex.

84. The bitstream of claim 83, wherein:

the flag allows a visual elementary stream of the bitstream as a standalone entity to specify a buffer model.

85. The bitstream of claim 83, wherein:

the at least one field whose inclusion in the VOL header is controlled by the flag designates a VOP rate of the bitstream.

86. The bitstream of claim 83, wherein:

the at least one field whose inclusion in the VOL header is controlled by the flag designates a peak bit rate of the bitstream.

87. The bitstream of claim 83, wherein:

the at least one field whose inclusion in the VOL

header is controlled by the flag designates whether the VOL contains at least one B-VOP.

88. The bitstream of claim 83, wherein:

the at least one field whose inclusion in the VOL header is controlled by the flag designates the size of the modeled buffer.

89. The bitstream of claim 83, wherein:

the at least one field whose inclusion in the VOL header is controlled by the flag designates said field that designates the occupancy of the buffer.

90. The bitstream of claim 82, wherein:

the video or visual object comprises one of a video object, a still texture object, a mesh object, or a face object.

91. The bitstream of claim 82, wherein:

the encoded data comprises a plurality of VOLs, and a decoder buffer model is applied independently to each VOL using buffer size and rate functions particular to each VOL.

92. The bitstream of claim 82, wherein:

the bitstream is compatible with an MPEG-4 video coding standard.

93. The bitstream of claim 82, wherein:

additional fields are provided in respective subsequent VOL headers to designate respective subsequent buffer occupancy levels.

94. The bitstream of claim 82, wherein:

said encoded data comprises video data that includes intraframe (I), prediction (P) and bi-directional (B) video object planes (VOPs).